Supplementary Materials

Resilience of Death: Intrinsic Disorder in Proteins Involved in the Programmed Cell Death

Zhenling Peng,[1] Bin Xue,[2] Lukasz Kurgan,[1,*] and Vladimir N. Uversky[2,3,4*]

[1]Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada;

[2]Department of Molecular Medicine, College of Medicine, University of South Florida, Tampa, FL 33612, USA;

[3]Byrd Alzheimer's Research Institute, College of Medicine, University of South Florida, Tampa, FL 33612, USA;

[4]Institute for Biological Instrumentation, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia

*To whom correspondence should be addressed: LK, Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta T6G 2V4, Canada; Phone: (780) 492-5488; Fax: (780) 492-1811; E-mail: lkurgan@ece.ualberta.ca; VNU, Department of Molecular Medicine, University of South Florida, 12901 Bruce B. Downs Blvd. MDC07, Tampa, Florida 33612, USA; Phone: 1-813-0748-5816; Fax: 1-813-974-7357; E-mail: vuversky@health.usf.edu

Materials and Methods

Datasets

In this work, several datasets of proteins related to the programmed cell death were analyzed. First, 1138 and 137 human proteins associated with apoptosis and autophagy, respectively, were collected from UniProt (1) on November 14, 2012. These proteins were selected using "reviewed:yes apoptosis human" and "reviewed:yes autophagy human" keywords and were grouped into human_apoptosis and human_autophagy sets, respectively. Since similar search for the human proteins associated with necroptosis gave only five hits, 35 human necroptosis-related proteins were manually picked based on the analysis of literature data (2-5). These proteins were also collected from the UniProt database and assembled into the human_necroptosis set.

The larger scale analysis of human proteins related to various forms of PCD that were collected from UniProt was next supplemented by a more focused analysis of fewer PCD-related proteins from several proteomes. Information about these proteins was collected from Deathbase (6), a specialized database dedicated to description of proteins involved in cell death, which includes higher-quality manually curated annotations. At this stage, we collected 3,458 proteins from Deathbase (6) on Nov 9th, 2011. This set includes proteins from 5 manually curated species: human, mouse, zebrafish, fly and worm, and 23 reference species that were annotated based on the similarity to the manually curated proteins; see Table S1. The proteins from the curated species include annotation of the corresponding cell death processes. They include 154, 11, 25, and 26 proteins that are annotated to participate in apoptosis, necroptosis, immune response-related cell death process, and other cell death processes, respectively.

Computational characterization of disorder

Proteins in the human_apoptosis, human_autophagy, and human_necroptosis sets were analyzed using PONDR-FIT algorithm (7), which is a meta-predictor that combines six individual predictors, which are PONDR® VLXT (8), PONDR® VSL2 (9), PONDR® VL3 (10), FoldIndex (11), IUPred (12), TopIDP (13). This meta-predictor is moderately more accurate than each of the component predictors and provides accurate disorder predictions at the residue level. The residue-level predictions allow for a more insightful analysis, including an investigation into the number and size of the predicted disordered segments. In addition to PONDR-FIT, two binary disorder classifiers, charge-hydropathy (CH) plot (14, 15) and cumulative distribution function (CDF) plot (15, 16), as well as their combination known as CH-CDF analysis (16-18), were used.

The primary difference between these two binary predictors (i.e., predictors which evaluate the predisposition of a given protein to be ordered or disordered as a whole) is that the CH-plot is a linear classifier that takes into account only two parameters of the particular sequence (charge and hydropathy), whereas CDF analysis is dependent on the output of the PONDR® predictor, a

nonlinear classifier, which was trained to distinguish order and disorder based on a significantly larger feature space. According to these methodological differences, CH-plot analysis is predisposed to discriminate proteins with substantial amount of extended disorder (random coils and pre-molten globules) from proteins with compact conformations (molten globule-like and rigid well-structured proteins). On the other hand, PONDR-based CDF analysis may discriminate all disordered conformations, including molten globules and mixed proteins containing both disordered and ordered regions, from rigid well-folded proteins. Therefore, this discrepancy in the disorder prediction by CDF and CH-plot provides a computational tool to discriminate proteins with extended disorder from potential molten globules and mixed proteins.

Positive and negative Y values in Figure 2D correspond to proteins predicted within CH-plot analysis to be natively unfolded or compact, respectively. On the other hand, positive and negative X values are attributed to proteins predicted within the CDF analysis to be ordered or intrinsically disordered, respectively. Thus, the resultant quadrants of CDF-CH phase space should be interpreted as follows: Q1, proteins predicted to be disordered by CH-plots, but ordered by CDFs; Q2, ordered proteins; Q3, proteins predicted to be disordered by CDFs, but compact by CH-plots (i.e., putative molten globules or mixed proteins); Q4, proteins predicted to be disordered by both methods (i.e., proteins with extended disorder).

Amino acid composition analysis of proteins in the human_apoptosis, human_autophagy, and human_necroptosis datasets was carried out using Composition Profiler (19) (http://www.cprofiler.org) using the PDB Select 25 (20) and the DisProt (21) datasets as reference for ordered and disordered proteins, respectively. Enrichment or depletion in each amino acid type was expressed as $(C_x\text{-}C_{order})/C_{order}$, i.e., the normalized excess of a given residue's content in a query dataset (C_x) relative to the corresponding value in the dataset of ordered proteins (C_{order}).

The disorder in the Deathbase proteins was predicted with MFDp method (22), which is a consensus-based predictor that was recently shown to provide accurate and competitive predictive quality (23, 24). MFDp predictions were used to calculate the disorder content (fraction of disordered residues), the number of disordered segments, and the number of long disordered segments that consists of at least 30 consecutive disordered amino acids; such long segments were found to be implicated in protein-protein recognition (25). We only counted the disordered segments with at least 4 consecutive disordered residues, which is consistent with (23, 26).

Function of the disordered segments was predicted based on a local pairwise alignment against functionally annotated disordered segments collected from DisProt 5.9 (21). We aligned each of the 10952 disordered segments extracted from the Deathbase into a set of 775 disordered segments collected from DisProt database that have functional annotations. We calculated alignment using the Smith-Waterman algorithm (27) based on the EMBOSS Water implementation with default parameters (gap_open=10, gap_extend=0.5, and blosum62 matrix). We defined sequence similarity as the number of identical residues in the local

alignment divided by the length of the local alignment or the length of the shorter of the two being aligned segments, whichever is larger. We transferred the annotation if the similarity is greater than 80%; this means that some of the segments could be annotated with multiple functions. Consequently, we successfully annotated 2108 disordered segments with 26 functions that are listed in Table S2. These annotations are used to investigate differences in the functional roles between short and long disordered segments extracted from the Deathbase. By considering all 128 annotated disordered segments from manually curated protein species, we also investigated whether the disordered segments involved in different cell death processes are associated with different functions.

We used MoRFpred (28), which is a leading predictor of molecular recognition features (MoRF), to annotate MoRF regions. MoRFs are short (5 to 25 amino acids) disordered regions that undergo disorder-to-order transition upon binding to protein partners and are implicated in signaling and regulatory functions (29-32). Following Mohan et al. (31), we grouped MoRF regions into α-MoRFs (that fold into α-helices), β-MoRFs (that fold into β-strands), γ-MoRFs (coils) and complex-MoRFs (mixture of different secondary structure), based on the secondary structure predicted with PSI-PRED (33).

We report sequence conservation for the ordered and disordered residues in the entire database and for each cell death process. The conservation was quantified with relative entropy (34) that was calculated from the Weighted Observed Percentages (WOP) profiles generated by PSI-BLAST (35). PSI-BLAST was run with default parameters (-j 3, -h 0.001) against the nr database, which was filtered using PFILT (36) to remove low-complexity regions, trans-membrane regions and coiled-coil regions. The use of the relative entropy is motivated by work in (34) that suggests that it leads to more biologically relevant results compared to some other conservation scores and the fact that it was recently applied to investigate disorder in histones (37) and in other related studies including identification of nucleotide-binding residues (38) and catalytic sites (39).

References

1. Consortium U. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res* 2012 Jan; **40** (Database issue)**:** D71-75.
2. Declercq W, Van Herreweghe F, Berghe TV, Vandenabeele P. Death receptor-induced necroptosis. *eLS*. John Wiley & Sons Ltd: Chichester, 2009.
3. Vandenabeele P, Galluzzi L, Vanden Berghe T, Kroemer G. Molecular mechanisms of necroptosis: an ordered cellular explosion. *Nat Rev Mol Cell Biol* 2010 Oct; **11** (10)**:** 700-714.
4. Bialik S, Zalckvar E, Ber Y, Rubinstein AD, Kimchi A. Systems biology analysis of programmed cell death. *Trends Biochem Sci* 2010 Oct; **35** (10)**:** 556-564.
5. Galluzzi L, Vanden Berghe T, Vanlangenakker N, Buettner S, Eisenberg T, Vandenabeele P*, et al.* Programmed necrosis from molecules to health and disease. *Int Rev Cell Mol Biol* 2011; **289:** 1-35.

6. Diez J, Walter D, Munoz-Pinedo C, Gabaldon T. DeathBase: a database on structure, evolution and function of proteins involved in apoptosis and other forms of cell death. *Cell Death Differ* 2010 May; **17** (5)**:** 735-736.
7. Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim Biophys Acta* 2010 Apr; **1804** (4)**:** 996-1010.
8. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK. Sequence complexity of disordered protein. *Proteins* 2001 Jan 1; **42** (1)**:** 38-48.
9. Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK, Obradovic Z. Optimizing long intrinsic disorder predictors with protein evolutionary information. *J Bioinform Comput Biol* 2005 Feb; **3** (1)**:** 35-60.
10. Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z. Length-dependent prediction of protein intrinsic disorder. *Bmc Bioinformatics* 2006; **7:** 208.
11. Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O, Beckmann JS*, et al.* FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* 2005 Aug 15; **21** (16)**:** 3435-3438.
12. Dosztanyi Z, Csizmok V, Tompa P, Simon I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 2005 Aug 15; **21** (16)**:** 3433-3434.
13. Campen A, Williams RM, Brown CJ, Meng J, Uversky VN, Dunker AK. TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. *Protein Pept Lett* 2008; **15** (9)**:** 956-963.
14. Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 2000 Nov 15; **41** (3)**:** 415-427.
15. Oldfield CJ, Cheng Y, Cortese MS, Brown CJ, Uversky VN, Dunker AK. Comparing and combining predictors of mostly disordered proteins. *Biochemistry* 2005 Feb 15; **44** (6)**:** 1989-2000.
16. Xue B, Oldfield CJ, Dunker AK, Uversky VN. CDF it all: consensus prediction of intrinsically disordered proteins based on various cumulative distribution functions. *FEBS Lett* 2009 May 6; **583** (9)**:** 1469-1474.
17. Mohan A, Sullivan WJ, Jr., Radivojac P, Dunker AK, Uversky VN. Intrinsic disorder in pathogenic and non-pathogenic microbes: discovering and analyzing the unfoldomes of early-branching eukaryotes. *Mol Biosyst* 2008 Apr; **4** (4)**:** 328-340.
18. Huang F, Oldfield C, Meng J, Hsu WL, Xue B, Uversky VN*, et al.* Subclassifying disordered proteins by the CH-CDF plot method. *Pac Symp Biocomput* 2012**:** 128-139.
19. Vacic V, Uversky VN, Dunker AK, Lonardi S. Composition Profiler: a tool for discovery and visualization of amino acid composition differences. *Bmc Bioinformatics* 2007; **8:** 211.
20. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H*, et al.* The Protein Data Bank. *Nucleic Acids Res* 2000 Jan 1; **28** (1)**:** 235-242.
21. Sickmeier M, Hamilton JA, LeGall T, Vacic V, Cortese MS, Tantos A*, et al.* DisProt: the Database of Disordered Proteins. *Nucleic Acids Res* 2007 Jan; **35** (Database issue)**:** D786-793.
22. Mizianty MJ, Stach W, Chen K, Kedarisetti KD, Disfani FM, Kurgan L. Improved sequence-based prediction of disordered regions with multilayer fusion of multiple information sources. *Bioinformatics* 2010 Sep 15; **26** (18)**:** i489-496.
23. Monastyrskyy B, Fidelis K, Moult J, Tramontano A, Kryshtafovych A. Evaluation of disorder predictions in CASP9. *Proteins* 2011; **79 Suppl 10:** 107-118.
24. Peng ZL, Kurgan L. Comprehensive comparative assessment of in-silico predictors of disordered regions. *Curr Protein Pept Sci* 2011 Oct 25.

25. Tompa P, Fuxreiter M, Oldfield CJ, Simon I, Dunker AK, Uversky VN. Close encounters of the third kind: disordered domains and the interactions of proteins. *Bioessays* 2009 Mar; **31** (3)**:** 328-335.
26. Noivirt-Brik O, Prilusky J, Sussman JL. Assessment of disorder predictions in CASP8. *Proteins* 2009; **77 Suppl 9:** 210-216.
27. Smith TF, Waterman MS. Identification of common molecular subsequences. *Journal of Molecular Biology* 1981 Mar 25; **147** (1)**:** 195-197.
28. Disfani FM, Hsu WL, Mizianty MJ, Oldfield CJ, Xue B, Dunker AK*, et al.* MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins. *Bioinformatics* 2012 Jun 15; **28** (12)**:** i75-83.
29. Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN, Dunker AK. Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry* 2005 Sep 20; **44** (37)**:** 12454-12470.
30. Vacic V, Oldfield CJ, Mohan A, Radivojac P, Cortese MS, Uversky VN*, et al.* Analysis of molecular recognition feature complexes. *Biophys J* 2007 Jan**:** 530a-530a.
31. Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, Dunker AK*, et al.* Analysis of molecular recognition features (MoRFs). *Journal of Molecular Biology* 2006 Oct 6; **362** (5)**:** 1043-1059.
32. Uversky VN, Dunker AK. Understanding protein non-folding. *Bba-Proteins Proteom* 2010 Jun; **1804** (6)**:** 1231-1264.
33. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology* 1999 Sep 17; **292** (2)**:** 195-202.
34. Wang K, Samudrala R. Incorporating background frequency improves entropy-based residue conservation measures. *Bmc Bioinformatics* 2006 Aug 17; **7**.
35. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W*, et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 1997 Sep 1; **25** (17)**:** 3389-3402.
36. Jones DT, Swindells MB. Getting the most from PSI-BLAST. *Trends Biochem Sci* 2002 Mar; **27** (3)**:** 161-164.
37. Peng Z, Mizianty MJ, Xue B, Kurgan L, Uversky VN. More than just tails: intrinsic disorder in histone proteins. *Mol Biosyst* 2012 Jul 6; **8** (7)**:** 1886-1901.
38. Chen K, Mizianty MJ, Kurgan L. Prediction and analysis of nucleotide-binding residues using sequence and sequence-derived structural descriptors. *Bioinformatics* 2012 Feb 1; **28** (3)**:** 331-341.
39. Johansson F, Toh H. A comparative study of conservation and variation scores. *Bmc Bioinformatics* 2010 Jul 21; **11**.
40. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P*, et al.* The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* 2011 Jan; **39** (Database issue)**:** D561-568.

Supplementary Tables

Table S1. Summary of the data collected from Deathbase. The curated and reference species are sorted by the size of their protein sets. The annotations of cell death processed are not available for the reference species.

	species	keyword	number of proteins					
			apoptosis	necroptosis	immune	other cell death	undefined	total
Manually curated species (228 proteins)	Human	H_sapiens	60	4	14	9	1	86
	Mouse	M_musculus	39	7	7	5	3	62
	Fly	D_melanogaster	18	0	4	12	0	35
	Zebrafish	D_rerio	27	0	0	0	7	34
	Worm	C_elegans	10	0	0	0	1	11
Reference species (3230 proteins)	Rat	R_norvegicus	n/a	n/a	n/a	n/a	n/a	249
	Chimpanzee	P_troglodytes	n/a	n/a	n/a	n/a	n/a	193
	Macaca	M_mulatta	n/a	n/a	n/a	n/a	n/a	192
	Dog	C_familiaris	n/a	n/a	n/a	n/a	n/a	183
	Cow	B_taurus	n/a	n/a	n/a	n/a	n/a	182
	Orangutan	P_pygmaeus	n/a	n/a	n/a	n/a	n/a	180
	Horse	E_caballus	n/a	n/a	n/a	n/a	n/a	174
	Fugu	T_rubripes	n/a	n/a	n/a	n/a	n/a	172
	Monodelphis	M_domestica	n/a	n/a	n/a	n/a	n/a	160
	Gasterosteus	G_aculeatus	n/a	n/a	n/a	n/a	n/a	159
	Medaka	O_latipes	n/a	n/a	n/a	n/a	n/a	152
	Tetraodon	T_nigroviridis	n/a	n/a	n/a	n/a	n/a	144
	Gorilla	G_gorilla	n/a	n/a	n/a	n/a	n/a	133
	Chicken	G_gallus	n/a	n/a	n/a	n/a	n/a	131
	Zebra finch	T_guttata	n/a	n/a	n/a	n/a	n/a	128
	Xenopus	X_tropicalis	n/a	n/a	n/a	n/a	n/a	126
	Rabbit	O_cuniculus	n/a	n/a	n/a	n/a	n/a	125
	Lyzard	A_carolinensis	n/a	n/a	n/a	n/a	n/a	124
	Ornithorhynchus	O_anatinus	n/a	n/a	n/a	n/a	n/a	106
	Anopheles	A_gambiae	n/a	n/a	n/a	n/a	n/a	69
	Ciona	C_intestinalis	n/a	n/a	n/a	n/a	n/a	68
	Aedes	A_aegypti	n/a	n/a	n/a	n/a	n/a	65
	Yeast	S_cerevisiae	n/a	n/a	n/a	n/a	n/a	15

Table S2. List of functional annotations together with the corresponding number of disordered segments for the 2108 disordered segments extracted from the Deathbase. Total of 26 functional annotations, where "protein-RNA binding" and "modification sites" combine several sub-function, were considered. The functions are sorted in the descending order by the total number of disordered segments that are divided into short, 4 to 30 amino acids (AAs), and long, 30 or more AAs, segments. The functions in bold font have less than 20 annotations or are not essential for proteins function and were excluded in the analysis shown in Figure 4B. The functions in either italics or bold have less than 5 annotations for the five curated species and were excluded in the analysis shown in Figure 4C. The counts of disordered segments given in italics correspond to individual cell death processes that were annotated for the five curated species.

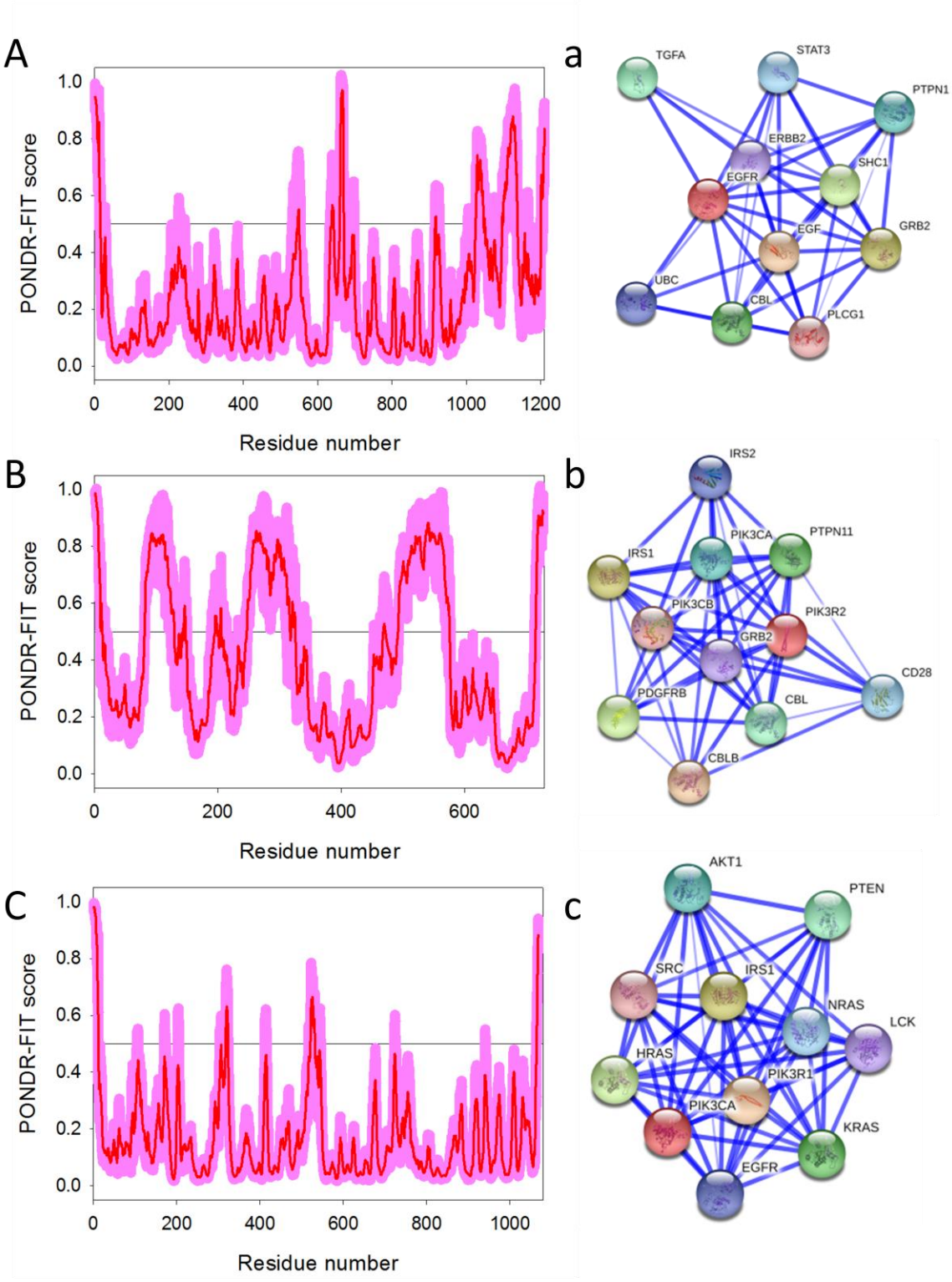
Function	Sub-function	# short (4 to 30 AAs) disordered segments	# long (≥30 AAs) disordered segments	# disordered segments (≥4 AAs) in apoptosis	# disordered segments (≥ 4 AAs) in other cell death processes	# disordered segments (≥ 4AAs) in immune responses	# disordered segments (≥ 4AAs) in necroptosis
Protein-protein binding		891	503	*56*	*14*	*6*	*12*
Substrate/ligand binding		331	158	*16*	*3*	*1*	*5*
Post-translational modification site	Phosphorylation	306	62	*12*	*2*	*2*	*4*
	Acetylation	37	0	*0*	*1*	*0*	*0*
	Fatty acylation	36	0	*2*	*0*	*0*	*0*
	Glycosylation	15	0	*0*	*0*	*0*	*0*
	Methylation	1	0	*0*	*0*	*0*	*0*
Protein-DNA binding	Protein-DNA binding	320	108	*13*	*4*	*0*	*1*
	DNA bending	1	0	*0*	*0*	*0*	*0*
	DNA unwinding	3	0	*0*	*0*	*0*	*0*
Flexible linkers/spacers		166	177	*8*	*5*	*1*	*1*
Intra-protein interaction		133	92	*7*	*1*	*0*	*2*
Protein-RNA binding	Protein-rRNA binding	74	0	*0*	*1*	*0*	*0*
	Protein-tRNA binding	51	0	*1*	*1*	*0*	*0*
	Protein-genomic RNA binding	44	0	*1*	*0*	*0*	*0*
	Protein-mRNA binding	27	0	*1*	*0*	*0*	*0*
	Protein-RNA binding	19	0	*0*	*0*	*0*	*0*
Electron transfer		63	101	*10*	*3*	*3*	*1*
Transactivation		145	18	*4*	*1*	*0*	*0*
Metal binding		133	25	*3*	*1*	*0*	*2*
Cofactor/heme binding		127	0	*5*	*1*	*0*	*0*
Protein-lipid interaction		59	66	*5*	*1*	*0*	*0*
Autoregulatory		108	11	*4*	*1*	*0*	*0*
Nuclear localization		111	0	*1*	*0*	*1*	*1*
Apoptosis Regulation		85	4	*5*	*1*	*1*	*0*
Entropic bristle		55	0	*1*	*0*	*0*	*0*
Protein inhibitor		37	0	*0*	*0*	*1*	*0*
Regulation of proteolysis in vivo		25	6	*0*	*1*	*0*	*0*
Polymerization		31	0	*2*	*0*	*0*	*1*
Self-transport through channel		19	0	*1*	*0*	*0*	*0*
Entropic clock		18	1	*0*	*0*	*0*	*0*
Entropic spring		6	0	*0*	*0*	*0*	*0*
Protein detergent		5	0	*0*	*0*	*0*	*0*
Structural mortar		5	0	*0*	*0*	*0*	*0*
Protein-Biocrystal binding		3	0	*0*	*0*	*0*	*0*
Sulfation		1	0	*0*	*0*	*0*	*0*
Not essential for protein function		32	75	*5*	*3*	*1*	*0*

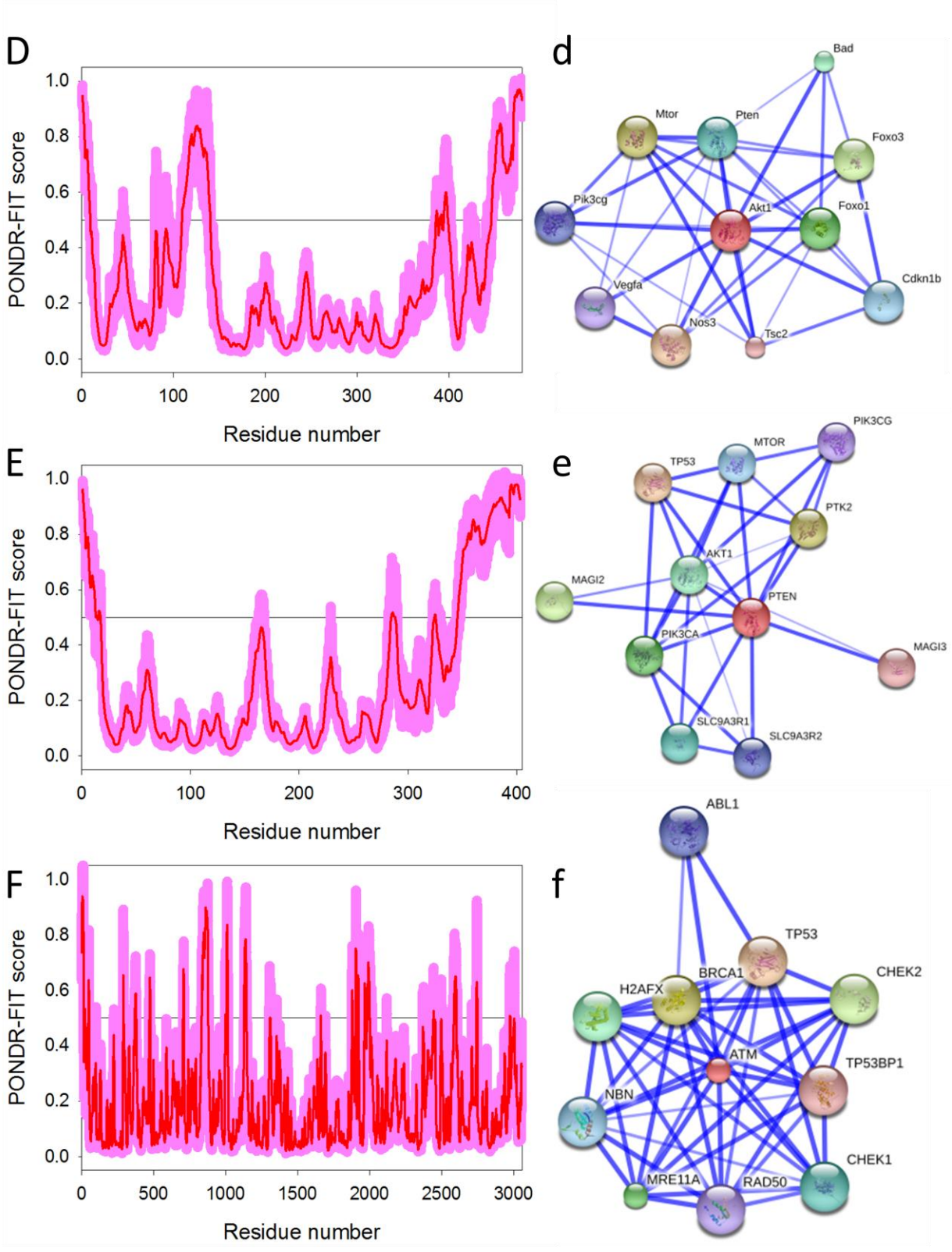
Supplementary Figures

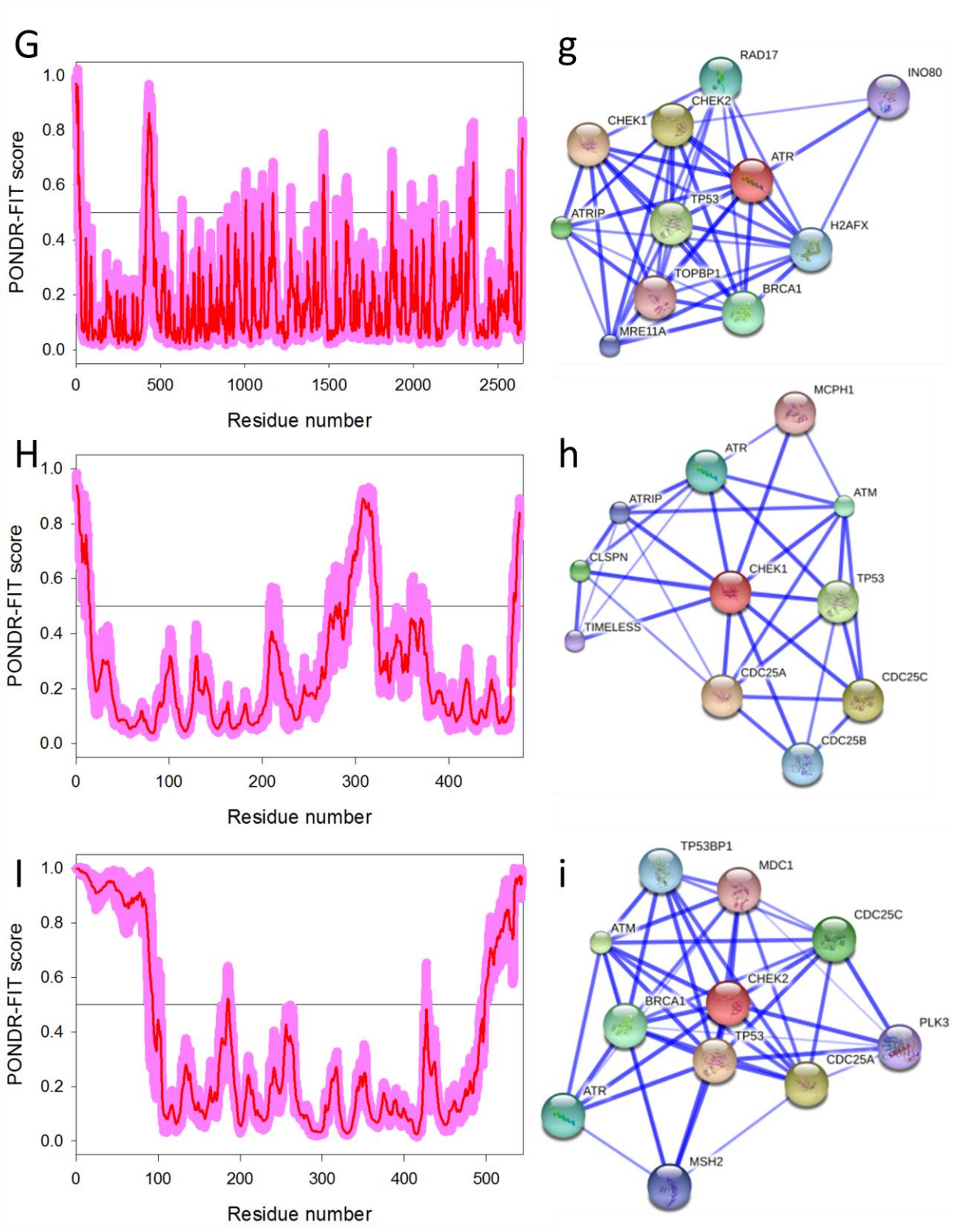
Figure S1. Intrinsic disorder (uppercase characters) and STRING analysis (lowercase characters) of the interactomes of some human proteins involved in the p53-mediated apoptotic signaling pathway.

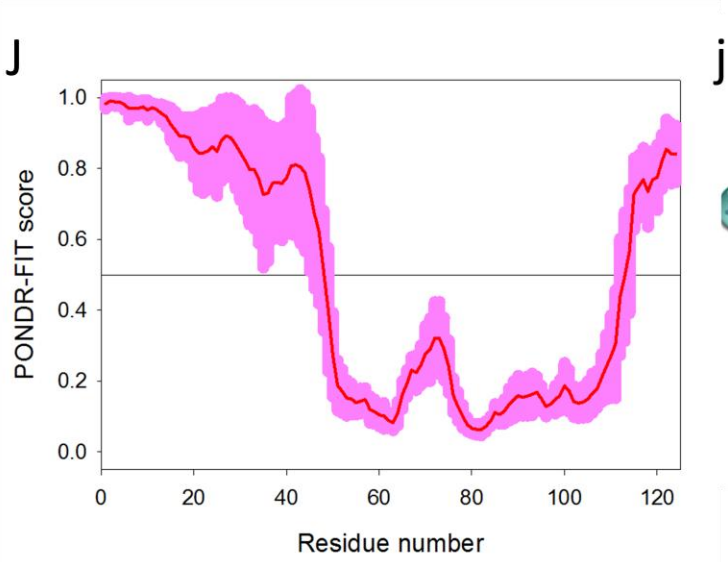
A. and **a**. EGFR (UniProt ID: P00553);
B. and **b.** PI3K regulatory subunit (UniProt ID: O00459);
C. and **c.** PI3K catalytic subunit (UniProt ID: P42336);
D. and **d.** AKT1 (UniProt ID: P47196);
E. and **e.** PTEN (UniProt ID: P60484);
F. and **f.** ATM (UniProt ID: Q13315);
G. and **g.** ATR (UniProt ID: Q13535);
H. and **h.** CHK1 (UniProt ID: O14757);
I. and **i.** CHK2 (UniProt ID: O96017);
J. and **j.** p53PAI-1 (UniProt ID: Q9HCN2);
K. and **k.** NOXA (UniProt ID: Q13794);
L. and **l.** BAX (UniProt ID: Q07812);
M. and **m.** BCL-2 (UniProt ID: P10415);
N. and **n.** BID (UniProt ID: P55957);
O. and **o.** Cytochrome *c* (UniProt ID: P99999);
P. and **p**. APAF-1 (UniProt ID: O14727);
Q. and **q.** Caspase-9 (UniProt ID: P55211);
R. and **r.** Caspase-3 (UniProt ID: P42574);
S. and **s.** Caspase-6 (UniProt ID: P55212);
T. and **t.** Caspase-7 (UniProt ID: P55210).

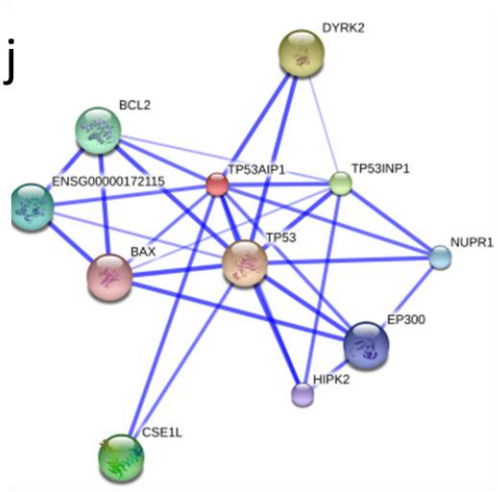
Intrinsic disorder propensity was evaluated by PONDR® FIT (red curves). Shadow around PONDR® FIT curves represents distribution of statistical errors. STRING database is the online database resource Search Tool for the Retrieval of Interacting Genes, which provides both experimental and predicted interaction information (40). For each protein, STRING produces the network of predicted associations for a particular group of proteins. The network nodes are proteins. The edges represent the predicted functional associations. The thickness of edges is proportional to the confidence level (40).

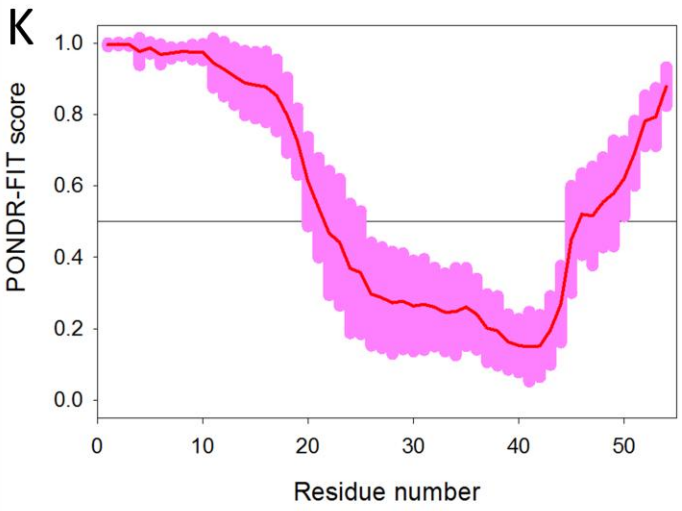

A
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
1200
Residue number
a
TGFA
STAT3
PTPN1
ERBB2
SHC1
EGFR
EGF
GRB2
UBC
CBL
PLCG1
B
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
Residue number
b
IRS2
PIK3CA
PTPN11
IRS1
PIK3CB
PIK3R2
GRB2
CD28
PDGFRB
CBL
CBLB
C
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
Residue number
c
AKT1
PTEN
SRC
IRS1
NRAS
LCK
HRAS
PIK3R1
PIK3CA
KRAS
EGFR

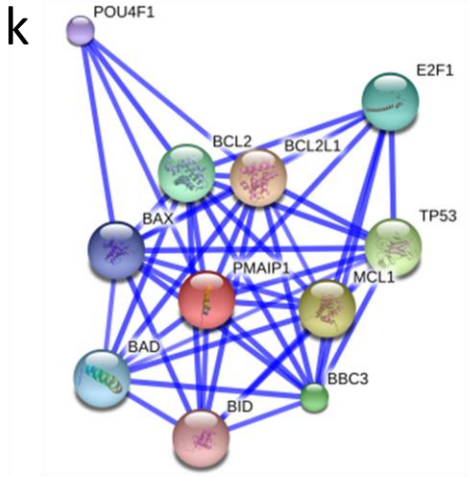

D
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
Residue number
d
Bad
Mtor
Pten
Foxo3
Pik3cg
Akt1
Foxo1
Cdkn1b
Vegfa
Nos3
Tsc2
E
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
Residue number
e
PIK3CG
MTOR
TP53
PTK2
AKT1
MAGI2
PTEN
PIK3CA
MAGI3
SLC9A3R1
SLC9A3R2
F
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
500
1000
1500
2000
2500
3000
Residue number
f
ABL1
TP53
BRCA1
CHEK2
H2AFX
ATM
TP53BP1
NBN
CHEK1
MRE11A
RAD50

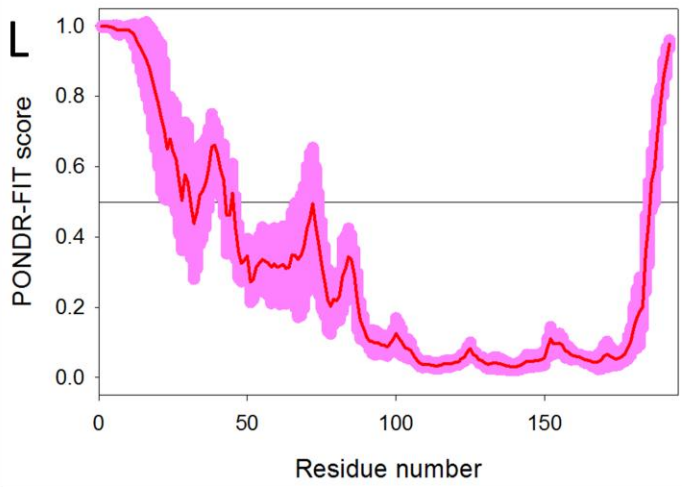

G
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
500
1000
1500
2000
2500
Residue number
g
RAD17
INO80
CHEK2
CHEK1
ATR
TP53
ATRIP
H2AFX
TOPBP1
BRCA1
MRE11A
H
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
Residue number
h
MCPH1
ATR
ATRIP
ATM
CLSPN
CHEK1
TP53
TIMELESS
CDC25A
CDC25C
CDC25B
I
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
500
Residue number
i
TP53BP1
MDC1
CDC25C
ATM
CHEK2
BRCA1
PLK3
TP53
CDC25A
ATR
MSH2

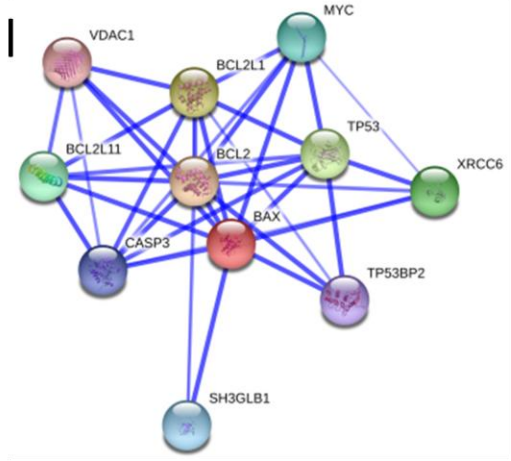

J
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
Residue number

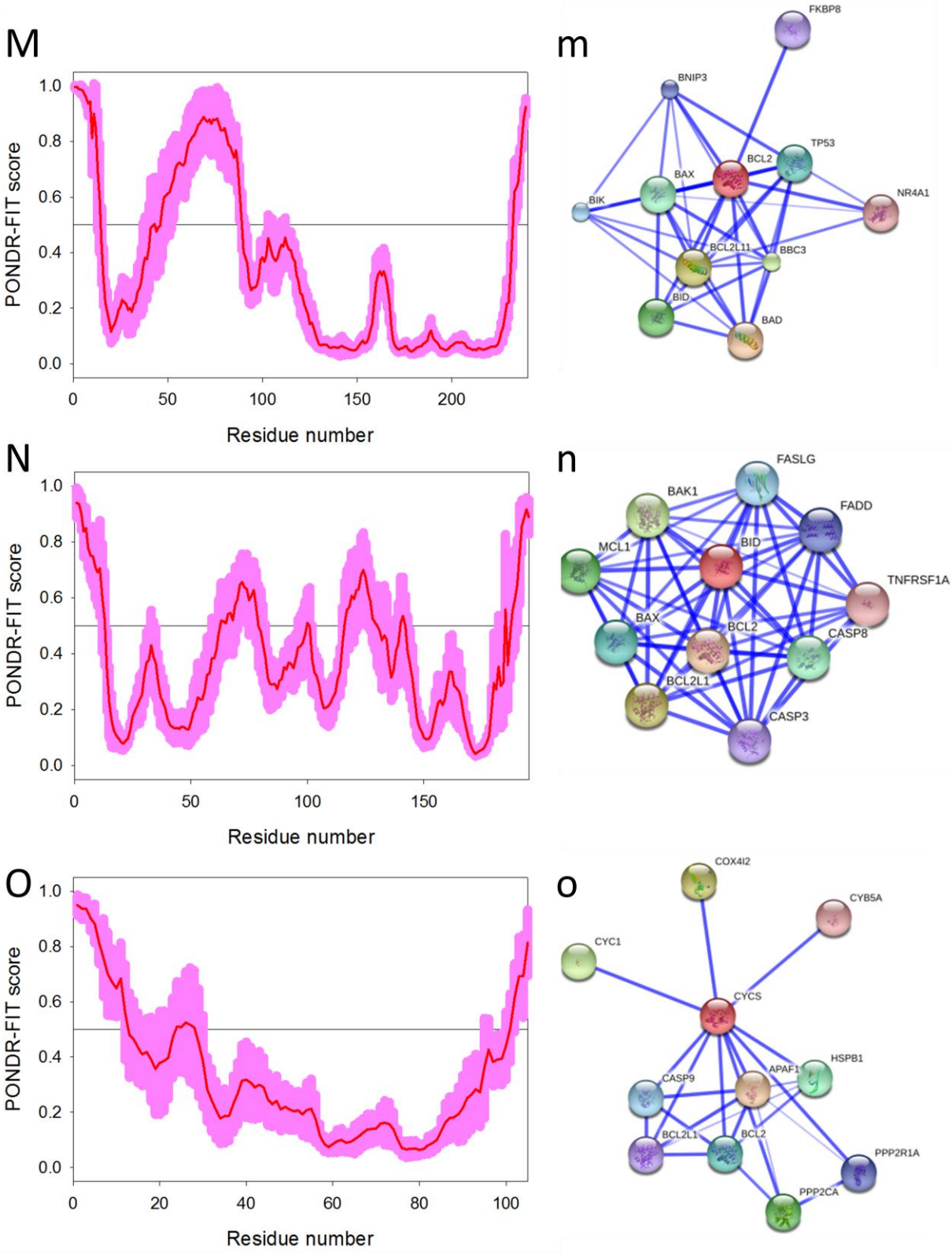

j
DYRK2
BCL2
TP53AIP1
TP53INP1
ENSG00000172115
TP53
NUPR1
BAX
EP300
HIPK2
CSE1L

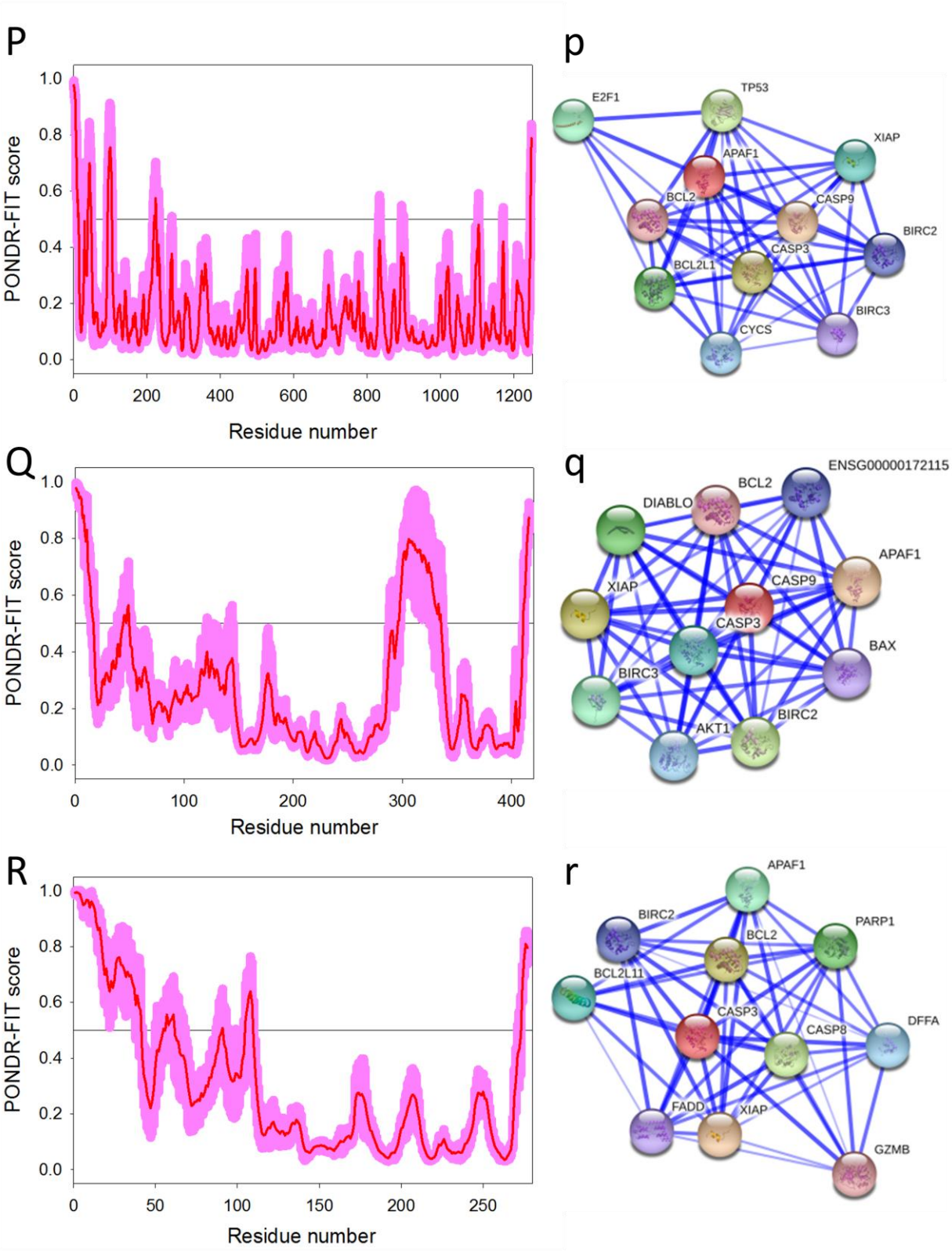

K
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
Residue number

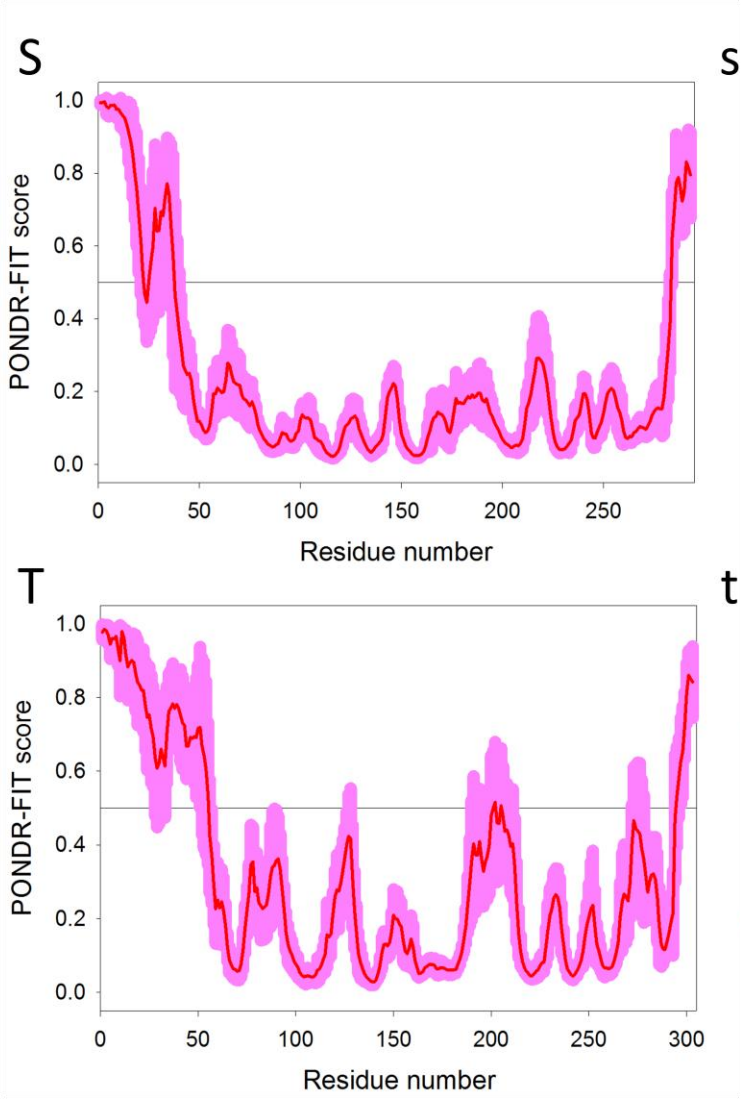

k
POU4F1
E2F1
BCL2
BCL2L1
BAX
TP53
PMAIP1
MCL1
BAD
BBC3
BID

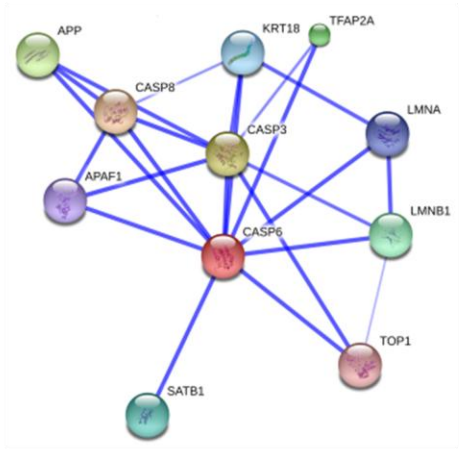

L
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
Residue number


l
MYC
VDAC1
BCL2L1
TP53
BCL2L11
BCL2
XRCC6
BAX
CASP3
TP53BP2
SH3GLB1

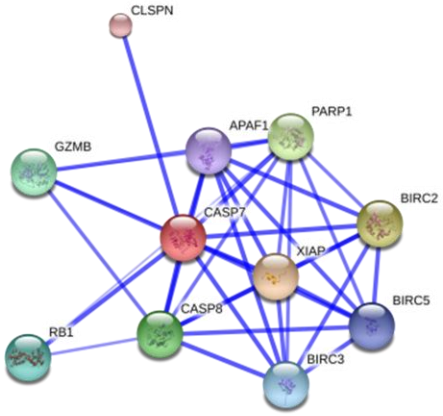

M
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
Residue number
m
FKBP8
BNIP3
TP53
BCL2
BAX
NR4A1
BIK
BCL2L11
BBC3
BID
BAD
N
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
Residue number
n
FASLG
BAK1
FADD
MCL1
BID
TNFRSF1A
BAX
BCL2
CASP8
BCL2L1
CASP3
O
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
Residue number
o
COX4I2
CYB5A
CYC1
CYCS
HSPB1
APAF1
CASP9
BCL2L1
BCL2
PPP2R1A
PPP2CA


P
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
1200
Residue number
p
E2F1
TP53
XIAP
APAF1
BCL2
CASP9
BIRC2
CASP3
BCL2L1
BIRC3
CYCS
Q
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
Residue number
q
ENSG00000172115
BCL2
DIABLO
APAF1
XIAP
CASP9
CASP3
BAX
BIRC3
BIRC2
AKT1
R
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
Residue number
r
APAF1
BIRC2
PARP1
BCL2
BCL2L11
CASP3
CASP8
DFFA
FADD
XIAP
GZMB


S
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
Residue number


s
APP
KRT18
TFAP2A
CASP8
LMNA
CASP3
APAF1
LMNB1
CASP6
TOP1
SATB1

T
PONDR-FIT score
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
300
Residue number


t
CLSPN
PARP1
APAF1
GZMB
BIRC2
CASP7
XIAP
BIRC5
CASP8
RB1
BIRC3